基于 DMD-LSTM 模型的股票价格时间序列预测研究 *

史建楠 [1]，邹俊忠 [1†]，张　见 [1]，汪春梅 [2]，卫作臣 [1]

(1. 华东理工大学 信息科学与工程学院，上海 200237; 2. 上海师范大学 信息与机电工程学院，上海 200234)

摘　要：针对股票市场关系复杂导致的有效特征提取困难、价格预测精度低等问题，提出一种基于动态模态分解-长短期记忆神经网络（DMD-LSTM）的股票价格时间序列预测方法。首先通过 DMD 算法对受市场板块联动效应影响的关联行业板块样本股数据进行分解计算，提取包含整体市场和特定股票走势变化信息的模态特征；然后针对不同市场背景，采用 LSTM 网络对基本面数据和模态特征进行价格建模预测。在鞍钢股份（sh600000）上的实验结果表明，该方法相较于传统预测方法，在特定的市场背景下能实现更高的价格预测精度，更为准确地描述股票价格的变化规律。

关键词：动态模态分解；长短期记忆神经网络；模态特征；板块联动效应；市场背景

中图分类号：TP391　　doi: 10.19734/j.issn.1001-3695.2018.08.0657

Research of stock price prediction based on dmd-lstm model

Shi Jiannan[1], Zou Junzhong[1†], Zhang Jian[1], Wang Chunmei[2], Wei Zuochen[1]

(1. *College of Information Science & Engineering, East China University of Science & Technology, Shanghai* 200237, *China*; 2. *College of Information Mechanical & Electrical Engineering, Shanghai Normal University, Shanghai* 200234, *China*)

Abstract: Aiming at the problems of low prediction accuracy and feature extraction difficulty in complicated stock market, this paper proposed a stock price prediction method based on dynamic mode decomposition and long short-term memory neural network (DMD-LSTM) . Firstly, it used the DMD algorithm to decompose the industry specific stock in the background of plate linkage phenomenon, and extracted the mode feature which included stock trend information. Then, built the LSTM network to establish the relations between stock price and the feature of mode and basic index in different market conditions. The experimental results on Angang Steel (sh600000) show that, the proposed method has the higher forecasting precision compared with the traditional ways in specific condition, which can characterize the trend of stock price changes better .

Key words: dynamic mode decomposition ; long short-term memory neural network; mode feature; plate linkage phenomenon; market condition

0　引言

股票市场作为国家经济的晴雨表，已经成为国家经济发展不可或缺的一部分。由于股票市场的重要性，通过分析挖掘股市历史数据和各项技术指标，并依据相关理论和算法模型来预测市场的价格走势情况具有重要的理论意义和社会价值[1]。股票市场作为复杂且时变的巨系统，其价格波动往往表现为较强的非线性[2]，因此如何从海量的市场数据中获取有价值的信息为决策服务是一个难点。

股票市场中，不同行业板块间的股票价格和收益率波动变化往往存在着一定的关联性，通常称之为联动效应。Bradly[3]采用方差分析法对同一证券市场不同板块股价联动效应进行深入研究，发现标普指数下不同行业指数间存在明显的联动效应；周武[4]在对国内 A 股市场多组行业板块进行分析后发现，不同市场经济周期下行业板块联动效应存在明显差异；魏国林[5]利用数据挖掘技术分析得出我国股市行业板块间的联动效应高于板块内各股联动效应。而在国内股市发展历程中，很多行情的大幅变化通常由多个板块带动发展而来[6]。因此借助相关行业板块趋势变化来发掘潜在关联信息，在股票价格预测上具有重要的研究意义。

由于股票市场数据指标众多且存在高维非线性的特性，目前的价格股票预测方法主要是将数据特征提取和模型预测过程相结合来进行研究。比如 Xu 等人[7]采取改进的经验模态分解算法（EEMD）对股票时间序列进行模态分解，并根据分解得到的模态信息成功预测市场价格趋势；Mann [8]将动态模态算法（DMD）用于证券市场这个高维非线性系统的信息提取上，并利用时空相干性结构信息来表征系统动态特征，准确的描述了市场价格变化和形态走势；李晋[9]利用离散小波分析来挖掘期货价格序列中隐含的趋势信息，然后引入支持向量机对重构的小波系数进行训练，大大提高了预测精度；郭志强[10]利用多线性主成分分析法（MPCA）进行特征提取和数据降维，并结合 RBF 神经网络在日经指数上进行了建模研究；史文静[11]对股指期货价格序列采用 EMD 算法得到 IMF 本征模函数，并结合 RBF 神经网络对提取出高低频 IMF 进

收稿日期：2018-08-31；**修回日期：**2018-10-23　**基金项目：**国家自然科学基金资助项目（61071085）

作者简介：史建楠（1993-），男，浙江象山人，硕士研究生，主要研究方向为模式识别与智能系统；邹俊忠（1960-），男，重庆人，教授，博士，主要研究方向为数据挖掘、机器学习、脑电信号处理（jzhzou2015@sina.com）；张见（1978-），男，工程师，博士，主要研究方向为数据挖掘和人工智能；汪春梅（1960-），女，江苏人，副教授，博士，主要研究方向为通信信号处理与模式识别；卫作臣（1991-），男，辽宁人，博士，主要研究方向为模式识别.

行建模来实现价格预测。

当前的股票预测方法大多直接选取市场数据为研究对象，没有考虑板块联动效应下关联行业板块整体走势间的影响和潜在关联因素；在特征提取方面对数据进行降维和转换等操作，缺乏业务可解释性；预测模型通常选取以统计学理论为基础的时间序列分析方法和传统机器学习方法，其在非线性时间序列预测准确性和时序信息利用方面均存在不足。

针对以上问题，本文从关联行业板块存在的联动效应出发，首先针对不同算法分别进行数据集构造，并通过 DMD（动态模型分解）算法对关联行业板块下样本股数据进行分解计算，提取代表行业板块的市场整体趋势以及股票自身价格趋势信息的模态特征。后续利用 LSTM 神经网络对非线性时间序列良好的拟合预测能力，分别构建 LSTM 模型对不同市场背景下股票基本面数据和模态特征进行学习，最终建立基于 DMD-LSTM 模型的股票价格预测方法及其融合模型。通过将不同市场背景下的模型预测结果与传统的机器学习方法进行分析对比，检验模型的有效性。

1　算法原理

1.1　DMD 算法

DMD（dynamic mode decomposition）动态模态分解由 Schmid 提出，并应用于流体分析、复杂系统构建以及数据挖掘方面[12]。DMD 算法围绕 Koopman 算符发展而来，其创新性的结合了时间维度上的功率谱分析和空间维度上的主成分分析的特点，通过将时间序列数据分解为特征值和特征向量，将高维非线性系统近似简化为低维线性系统表示，最终实现系统当前状态重构和短期状态预测[13]。

设 $x_1,x_2,\cdots,x_M$ 为原始数据中 M 个时间序列截面数据，每个时间序列截面数据包含 N 个样本，记为 $N\times M$ 矩阵 $X=[x_1,x_2,\cdots,x_M]$，称为数据快照（data shot）。

根据 Koopman 算符线性映射的思想，总能找到一个矩阵 A 使得系统下一时刻的状态由上一时刻的状态进行表示，即 $x_{j+1}=Ax_j$。将数据快照数据表示为 Krylov 子序列形式

$$X_1^{M-1}=[x_1,x_2,....,x_{M-1}]=[x_1,Ax_1,...,A^{M-2}x_1] \tag{1}$$

$$X_2^{M-1}=[x_2,x_3,...,x_M]=AX_1^{M-1} \tag{2}$$

矩阵 A 为系统由 X_1^{M-1} 向 X_2^{M-1} 状态演变规律的线性估计，则 X_2^{M-1} 中 x_M 可以由上一个时刻 X_1^{M-1} 中线性独立的 x_i（列矢量）的加权线性组合来表示：

$$x_M=\sum_{m=1}^{M-1}a_m x_m+r=X_1^{M-1}a+r \tag{3}$$

其中：a 为加权线性组合的权值矢量，r 为长度为 N 的残差矢量。根据上式可得关系：

$$AX_1^{M-1}=X_2^{M-1}=X_1^{M-1}S+re_{M-1}^T \tag{4}$$

e_{M-1}^T 为单位矢量，通过极小化残差矩阵 re_{M-1}^T 来寻找矩阵 A 分解后最优的低维特征值和特征矢量，可以使得系统估计值和目标 x_M 最接近。S 为低维近似矩阵，矩阵 A 的特征值可由矩阵 S 的特征值来近似替代。

$$S=\begin{bmatrix} 0 & & & & a_1 \\ 1 & 0 & & & a_2 \\ & \ddots & \ddots & & \vdots \\ & & 1 & 0 & a_{M-2} \\ & & & 1 & a_{M-1} \end{bmatrix} \tag{5}$$

矩阵 S 最后一列矢量 a 的值可以表示为

$$a=(X_1^{M-1})^{-1}x_M \tag{6}$$

实际求解中矩阵 S 的性质通常较差不易求解，因此进一步采用 S 的相似矩阵 $\tilde{S}$ 来进行计算求解。对式（5）中矩阵 X_1^{M-1} 进行约化 SVD 分解，并进行转换可得

$$X_2^{M-1}=U\,\Sigma\,V^H S \tag{7}$$

$$S=V\,\Sigma^{-1}U^H X_2^{M-1} \tag{8}$$

$$\tilde{S}=U^H X_2^{M-1}V\,\Sigma^{-1}=U^H AU\equiv\tilde{A} \tag{9}$$

将计算矩阵 A 的特征值问题转换为计算矩阵 $\tilde{S}$ 的特征值 u_k 和特征矢量 y_k，特征值 u_k 包含着以 Δt 为时间间隔的系统向下一个时刻状态变化的时间动态特征

$$\tilde{S}y_k=u_k y_k \quad k=1,2,....,K \tag{10}$$

系统的 DMD 动态模态可以由矩阵 A 第 k 个特征矢量 ψ_k 表示，根据相似矩阵的性质可得

$$\psi_k=Uy_k \tag{11}$$

最终，利用 ψ_k 和 u_k 可以实现系统当前状态重构和将来状态预测。改写特征值形式为 $\omega_k=\ln(\mu_k)/\Delta t$，则系统 t 时刻截面数据可以表示为

$$x_{DMD}(t)=\sum_{k=1}^{K}b_k(0)\psi_k(x)\exp(\omega_k t)=\Psi diag(\exp(\omega t))\,b \tag{12}$$

其中：$b_k(0)$ 为模态初始幅值，Ψ 为特征向量 ψ_k 组成的矩阵，$diag(\omega t)$ 是对角元为 $\exp(\omega t)$ 的对角阵，b 为 $b_k(0)$ 构成的向量。向量 b 通过 0 时刻的截面数据 x_1 进行求解，利用 DMD 进行状态重构得到 x_1，然后求矩阵伪逆即可。

$$x_1=x_{DMD}(0)=\Psi diag(\exp(0))\,b=\Psi\cdot b \tag{13}$$

$$b=\Psi^{+}x_1 \tag{14}$$

DMD 算法通过提取低维动态特征来表征系统变化，实现以线性映射去最优拟合系统的非线性动态情况，最终利用提取的 DMD 动态模态实现对系统任意时刻的状态估计。

1.2　LSTM 神经网络

循环神经网络（RNN）是目前较为流行的应用于时间序列预测的神经网络模型。传统的前向神经网络比如 BPNN，神经元信号只能单向流动，导致对样本的处理在时间上是独立的，对时间序列数据而言损失了重要的时序信息[14]。RNN 在传统神经网络结构基础上进一步对隐层神经元之间也引入了权连接，使得隐层神经元每个时刻的输出都依赖于之前时刻的信息，网络单元既存在前馈连接又有内部反馈连接，实现了动态时序行为在神经网络单元内部状态的转换。

但在实际应用中，RNN 存在梯度消失和梯度爆炸导致的长期依赖问题，使得其无法应用于较长时间序列数据的建模。而由 Grves[15]改进提出的 LSTM（Long Short-Time Memory）模型作为 RNN 的一种完善成功解决了上述问题。LSTM 利用特定的记忆单元代替了 RNN 链式结构中的隐层节点，实现了时序信息保留和长期记忆的能力。图 1 为 LSTM 中的记忆单元结构。

LSTM 记忆单元的主要结构是门(gates)和记忆细胞（cell），门结构可以让信息选择性的通过实现信息的去除或者增加，实现记忆细胞状态的保持和更新，具体包括输入门（input gates）、输出门（output gates）、遗忘门（forget gates）；记忆细胞负责存储细胞状态信息。

遗忘门将当前时刻的输入值 $x^{(t)}$ 和上一时刻的隐藏状态值 $h^{(t-1)}$ 作为 sigmoid 函数的输入，产生一个[0,1]的权值输出 $f^{(t)}$。权值可看做记忆单元遗忘上一层细胞状态的概率，通过将 $f^{(t)}$ 乘以记忆单元上一刻的细胞状态 $C^{(t-1)}$ 来控制上一单元信息被保留程度。

$$f^{(t)}=\sigma(W_f\cdot[h^{(t-1)},x^{(t)}]+b_f) \tag{15}$$

输入门利用 sigmoid 函数和 tanh 函数配合来控制能够进入细胞状态的新输入信息的量。其中 tanh 函数生成新信息 $\tilde{C}^{(t)}$，sigmoid 函数产生[0,1]的权值来控制进入细胞状态的信息量大小。进一步将输入门和遗忘门相结合就可以实现将细

胞状态由 $C^{(t-1)}$ 更新为 $C^{(t)}$。

$$i^{(t)}=\sigma(W_i\cdot[h^{(t-1)},x^{(t)}]+b_i) \tag{16}$$

$$\tilde{C}^{(t)}=\tanh(W_a\cdot[h^{(t-1)},x^{(t)}]+b_a) \tag{17}$$

$$C^{(t)}=f^{(t)}*C^{(t-1)}+i^{(t)}*\tilde{C}^{(t)} \tag{18}$$

输出门通过对当前细胞状态 $C^{(t)}$ 使用 tanh 函数进行处理得到进入下一时刻的细胞状态信息，然后配合 sigmoid 函数得到的权值 $o^{(t)}$ 来控制细胞状态信息的过滤程度，并最终得到下一时刻的隐藏信息 $h^{(t)}$。

$$o^{(t)}=\sigma(W_o\cdot[h^{(t-1)},x^{(t)}]+b_o) \tag{19}$$

$$h^{(t)}=o^{(t)}*\tanh(C^{(t)}) \tag{20}$$

LSTM 通过更换 RNN 链式结构中记忆单元实现了对信息流的控制，较为成功的解决了 RNN 中出现的梯度问题，实现了长周期信息的记忆。目前 LSTM 已经在多种时间序列数据预测上获得了广泛的应用，包括 NLP 任务，证券市场价格预测等。

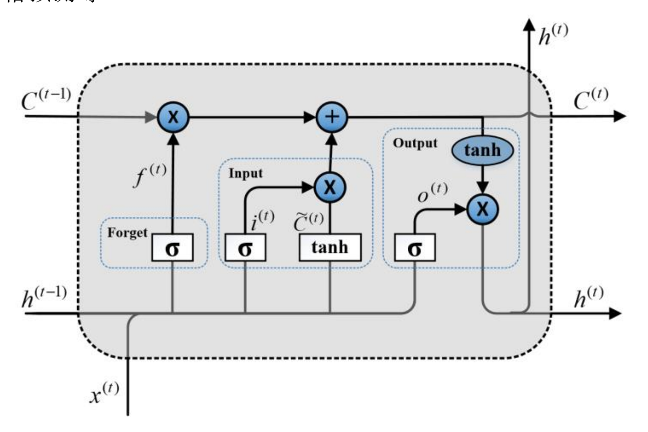


图 1　LSTM 记忆单元结构

Fig. 1　Structure of LSTM memory unit

2　基于 DMD 和 LSTM 的股票价格预测方法

2.1　模型数据集构造

在时间序列分析中，数据样本通常采用滑动时间窗技术来进行切分构建，本文针对 DMD 算法和 LSTM 模型的不同分别进行数据样本构建。

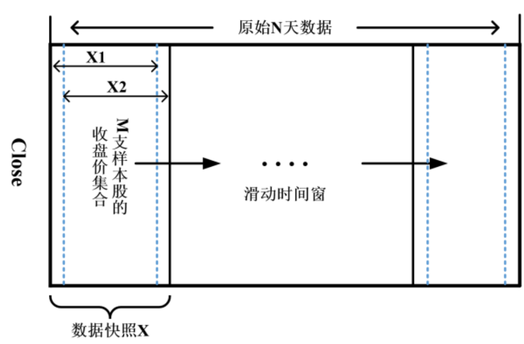


图 2　DMD 数据快照矩阵切分构建

Fig. 2　Construction of DMD Snapshot matrix

图 2 展示了 DMD 数据快照矩阵的构建过程。DMD 算法提取模态特征（mode feature）的源数据为关联行业板块下 M 支样本股的 N 天的收盘价数据集合，设定滑动步长为 1，通过滑动时间窗来切割源数据获取 DMD 算法所需的数据快照矩阵 X。为了消除不同股票价格间的差异性，对数据快照 X 中样本股数据分别进行归一化（minmax scaler）处理。

LSTM 模型进行建模的数据为股票的基本面特征（raw feature）和提取的 DMD 模态特征数据组成。设滑动窗口宽度为 L，采用滑动时间窗从数据集中截取数据单元构建样本，并取窗口后一天的收盘价为作为预测目标。通过不断后移，形成一系列相互覆盖的样本数据，在降低数据时变特性的同时，利用了数据的时序信息。

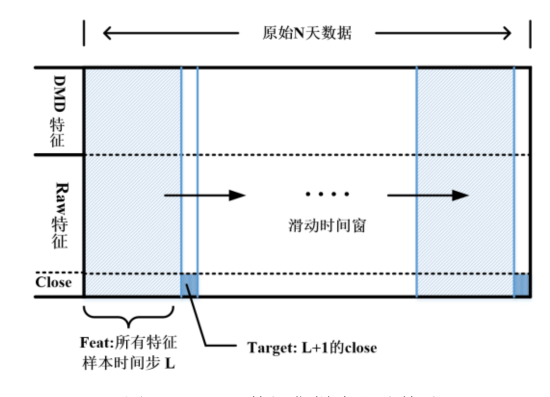


图 3　LSTM 数据集样本切分构建

Fig. 3　Construction of LSTM dataset

最终得到的数据集中不同特征值范围差异较大，不利于 LSTM 模型收敛，同样需要对 LSTM 数据集中特征分别进行归一化操作来调整数据分布。

2.2　基于 DMD 模态分解的特征提取

DMD 算法适用于复杂系统分析，但其应用于股票市场的前提假设是：股票市场是整体相关的复杂系统，即某个行业板块的股票价格波动可能受到相关行业板块的影响。在此基础上利用 DMD 分别提取代表过去 n 天的市场整体趋势变化和特定股票价格变化的模态特征，捕捉股票价格变化的动态特性和潜在行为模式。实证研究中 DMD 算法的应用对象来自于沪深股市中 5 个行业板块，其中主导行业为钢铁行业，关联行业为有色金属、金属制品、汽车制造业、建筑业。

利用 DMD 对数据快照矩阵 X 进行模态分解后可以获取整个系统的动态模态，以特征值 u_k 来表示。特征值 u_k 分解顺序对应着市场动态模态信息量的大小，将系统分解得到的第一个特征值 u_1 称为主导特征值，又称主导模态（domain mode），其代表整个市场的主要趋势信息。由于特征值一般为复数，Mann[8]指出当特征值位于在单位圆外，其模值大于 1，意味着市场有扩张趋势；当特征值位于在单位圆内，其模值小于 1，市场有紧缩趋势；当特征值位于在单位圆上时，表示市场整体趋势平稳。特别对于位于横轴没有虚部的特征值，对应模态下的市场趋势变化速率是指数形式的，因此该模态对于系统的影响相比于其他模态更加显著，是主导市场变化的重要因素之一。

图 4 为某个时间窗为 20 的数据快照矩阵 X 进行 DMD 模态分解运算的结果。左子图为分解得到的特征值 u_k 在复平面上的分布，对应于总体市场的模态特征；右子图为 u_k 经过傅里叶转换后的 ω_k 在复平面的分布。图中三角形点为单位圆外包含扩张趋势信息的特征值，对应于复平面中实部大于 0 的 ω_k。

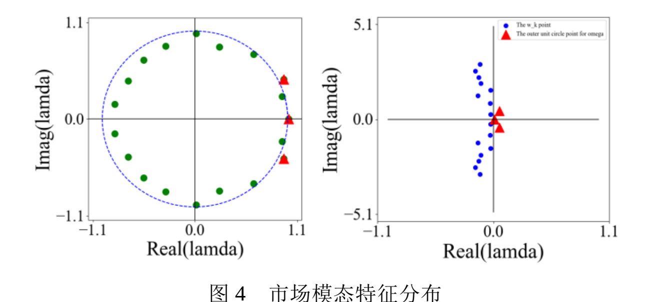


图 4　市场模态特征分布

Fig. 4　Distributions of market's mode feature

图 5 为选取的鞍钢股份（sh00898）经过 DMD 分解得到的前 10 个动态模态，来源于特征向量 ψ_k，对应价格趋势变化中不同层次的潜在信息。前三个主要模态值大于 0 呈现扩张趋势，后面影响因素较小的模态值小于 0 呈现下跌趋势。

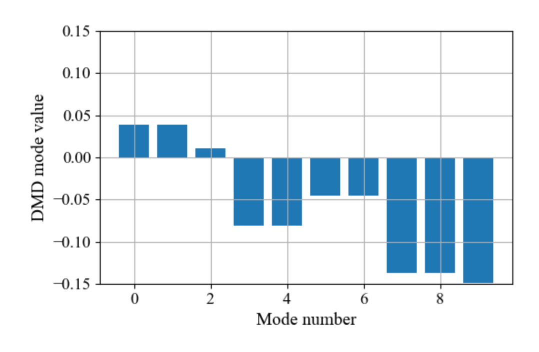


图 5　鞍钢股份分解得到的动态模态

Fig. 5　DMD mode value of decomposition's result(Angang Steel)

2.3　DMD-LSTM 价格预测模型

本文综合 DMD 算法和 LSTM 模型来进行股票价格预测，DMD 算法用来对特定板块下的股票集合的日收盘价数据快照进行分解计算来提取模态特征，LSTM 结合股票市场的基本面特征和提取的模态特征来进行模型训练，将股票收盘价的一步预测问题转换为监督学习问题。

股票市场形态主要分为单边形态、震荡形态、平稳形态和其他特定形态，不同的市场形态下数据变化情况不同，模型的预测效果存在差异。DMD 主要依据数据历史趋势来进行模态的提取，因此模态分解得到的结果存在一定的滞后性。当市场处于单边形态时，DMD 抽取的模态在市场趋势惯性下能较为准确的抓取市场变化的趋势信息；当市场处于震荡市场的时候，由于滞后性会使得 DMD 分解得到的模态对市场趋势变化的跟踪效果会减弱。因此本文分别就单边上涨、单边下跌、震荡三种不同的市场形态进行实证研究。

LSTM 对股票价格进行建模需要指定数据的时间步和合适的特征。设定 LSTM 模型的样本时间步 L=20，样本特征由两部分总计 9 维特征组成，其中基本面特征包括单只特定股票的交易日收盘价（close）、开盘价（open）、最高价（high）、最低价（low）、成交量（volume）5 维特征，DMD 模态特征为市场主导模态特征和对应股票的前 3 个动态模态特征共计 4 维特征。LSTM 模型的网络结构和参数设置决定着模型的性能，因此需要通过相关指标定量的评价网络模型预测性能来获取最优的网络结构配置，实现股票价格与特征关系最佳映射。

在确定 LSTM 模型结构后，基于已训练的模型对测试集样本进行预测。在 DMD-LSTM 单模型基础上，根据模型融合的思路，将 DMD-LSTM 模型和基于基本面特征的 LSTM 单模型的预测结果加权平均作为预测值。最后对预测价格进行反归一化转换为实际价格量纲，通过评价指标来判断模型预测性能，整体模型预测流程见图 6。

3　实证分析

3.1　数据来源

选取钢铁行业板块中的鞍钢股份(sh00898)为研究对象，同时为了配合 DMD 计算需求，选取钢铁行业及相关行业下共计 117 支股票的收盘价来构建数据快照。

鉴于市场形态的不同，本文分别就股票单边上升期，单边下降期和价格震荡期三种市场形态进行实证研究。源数据为鞍钢股份 2014-05-27 到 2018-05-27 期间的日交易数据，选取 2017-05-08 到 2017-08-02 的股票价格单边上涨期，共计 60 天数据作为验证集，对应选取 2014-05-27 到 2017-05-07 约 700 个交易日数据作为训练集；选取 2017-02-22 到 2017-05-22 的股票价格单边下跌期，共计 60 天数据作为验证集，对应选取 2017-05-27 到 2017-02-21 约 650 个交易日数据作为训练集；选取 2017-09-27 到 2017-12-27 的价格震荡期，共计 60 天数据作为验证集，对应选取 2014-05-27 到 2017-09-26 约 800 个交易日数据作为训练集。实验数据来源于新浪财经接口，数据集样本的划分构建和预处理见 2.1 节。

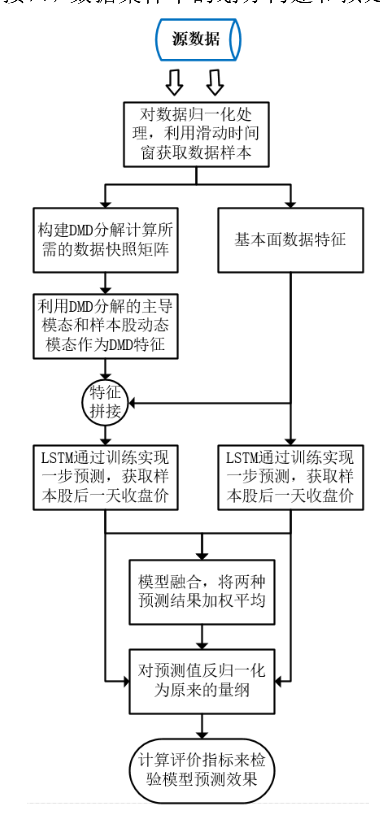


图 6　DMD-LSTM 模型进行股票价格预测流程图

Fig. 6　Prediction procedure of DMD-LSTM model

3.2　模型评价指标

在对模型预测效果进行定量评价时，本文对于价格预测能力的评价指标选取为均方根误差 RMSE、平均绝对误差 MAE、确定系数 R^2，其中 RMSE 和 MAE 越小，模型对于真实值拟合的偏差越小，结果越准确；确定系数 R^2 越接近 1，代表拟合优度越大，模型拟合数据能力效果越好。价格趋势变化预测能力的评价指标选取为方向准确率 DA，其定义为模型对股票价格变化方向预测的准确性，具体公式定义如下：

$$RMSE=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i-y_i)} \tag{21}$$

$$R^2=SSR/SST=\sum_{i=1}^{N}(\hat{y}_i-\bar{y})^2\Big/\sum_{i=1}^{N}(y_i-\bar{y})^2 \tag{22}$$

$$MAE=\frac{1}{N}\sum_{i=1}^{N}|(\hat{y}_i-y_i)| \tag{23}$$

$$DA=\frac{1}{N}\sum_{i=1}^{N}d_i \tag{24}$$

其中：N 为真实值和预测值对比次数，$\hat{y}_i$ 为模型预测值，y_i 为真实值，$\bar{y}$ 为 y_i 的均值，d_i 为预测方向标志位，方向预测正确 d_i 置 1，方向预测错误 d_i 置 0。

3.3　DMD-LSTM 模型预测方法实现

3.3.1 DMD 模态特征提取

DMD 通过模态的提取来描述价格的变化规律，其中市

场主导模态的变化代表着市场总体趋势随着时间变化的规律。如图 7，主导模态模值总体围绕 1 进行波动，对于模值偏离 1 的模态，代表市场中有较强的上涨或者下跌的动力。

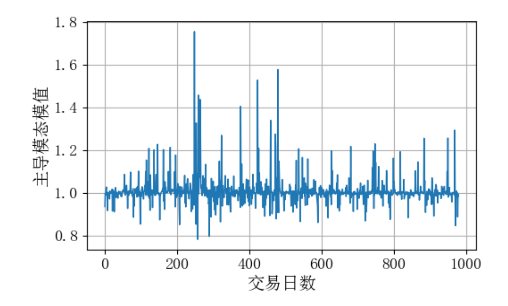


图 7 市场主导模态的时间序列曲线

Fig. 7 The curve of market's domain mode value

图 8 为鞍钢股份全量数据进行 DMD 模态分解后，前三个主要的动态模态的日变化曲线。通常股票价格中长期走势相对来说变化较缓，可看做低频信号；而短期价格走势波动较大，可看做高频信号。类似的，动态模态 1 可看做低频高能量模态，其波动范围较为稳定，表示该股票的价格变化趋势相对较慢。模态 2、3 波动相较于模态 1 有所增加，相比于模态 1 为高频低能量的模态，波动相对较大，说明短期价格波动会比较显著。

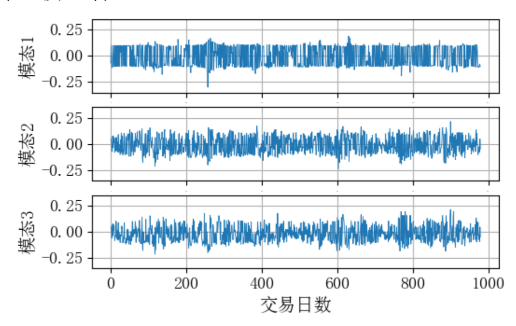


图 8 鞍钢股份前三个主要模态的变化曲线

Fig. 8 The curve of the first three mode values(Angang Steel)

3.3.2 LSTM 模型结构选定

神经网络的网络结构和参数设置决定着模型的性能。为了客观分析模型的真实预测能力，消除参数选取差异对实证结果的影响，本文通过模型参数调优实现股票价格与特征关系的最佳映射。

本文的 LSTM 模型结构由输入层、两层 LSTM 层和输出层组成，损失函数使用均方根误差，模型训练过程中采用 Adam 算法进行优化。模型参数选择上包括学习率、隐层神经元个数、训练次数和正则化参数，网络模型的搭建在 Tensorflow 框架下实现。

以股票价格单边上涨情况为例，LSTM 神经网络调参过程如下：首先固定训练次数 epoch=800，进行隐层神经元个数和学习率的调节，将两层 LSTM 层中的节点个数从 10 到 30 间隔为 2 进行选取，调节学习率以间隔为 0.0001 从 0.0001 到 0.001 进行选取，两者总计 100 组参数，确定最佳的的隐层节点数和学习率组合为（14，14，0.0005）。epoch 的大小会对模型的预测精度和运行效率造成较大影响，较大的 epoch 导致运行效率低并可能伴随过拟合现象，较小的 epoch 会使模型训练不充分出现欠拟合现象。图 9 为 LSTM 模型在训练过程中 loss 关于 epoch 的变化曲线，根据 loss 变化曲线确定最佳 epoch 为 140。为了增加模型泛化能力消除过拟合现象，本文在 LSTM 网络结构中增加 L2 正则项，正则项系数为 0.0001。

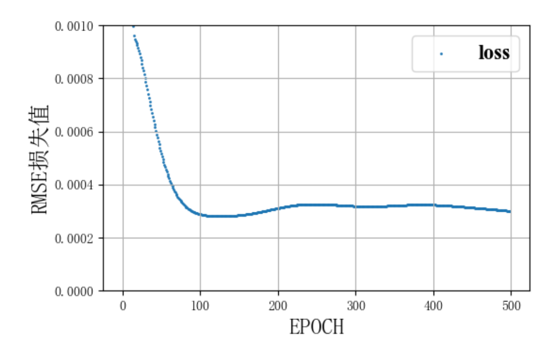


图 9 单边上涨形态下 LSTM 的 loss 变化曲线

Fig. 9 The loss in LSTM training of single market

参照上述参数配置过程，选取对比实验模型的最佳的模型参数，基于 RBF 核的支持向量回归机(SVR)的正则化系数和 gamma 参数组合为（4，0.1）；基于 SGD 优化的 BP 神经最佳的学习率、隐层层数、隐层节点数、epoch 的参数组合为（0.15，2，40，4500），DMD 滑动窗口参数为 20。对于股票价格单边下降和震荡形态，LSTM 模型参数配置过程相同。

3.3.3 结果对比与分析

为了检验模型的真实预测性能，消除偶然因素的影响，实证中针对不同的市场情况，分别对每种模型进行 10 次实验来取平均结果进行比较。

股票价格单边上涨阶段，市场趋势总体走强，相关模型的预测结果如表 1 所示。SVR 模型和 BP 模型作为传统的预测方法，实证中以两者的预测结果作为基准进行对比。DMD 作为无监督的数据挖掘方法，由结果可知其在价格预测上的精度较差，在 RMSE 指标上比基准水平低了平均 35.61%。由于市场趋势较为显著，DMD 在价格趋势方向的预测准确性相对较高，DA 指标比基准模型高了平均 22.60%，主要原因是 DMD 能利用时序信息从市场行为中发掘趋势变化的有效信息。DMD-LSTM 模型的价格预测精度高于其他单模型，包括仅利用基本面指标的 LSTM 模型，这说明 DMD 提取的模态特征一定程度上包含着市场行为变化的潜在信息，能提高 LSTM 模型的预测精度，其在 RMSE 指标上比基准水平高了平均 12.02%。在模型融合方面，将 LSTM 单模型和 DMD-LSTM 单模型的预测结果进行加权平均，价格预测精度比 DMD-LSTM 模型高了 1.34%，同时模型的模型拟合优度最佳和 DA 值也有所提升，实现了预测效果的最优。图 10 为选取的三个结果差异较大的模型 DMD、BP、融合模型的预测结果。

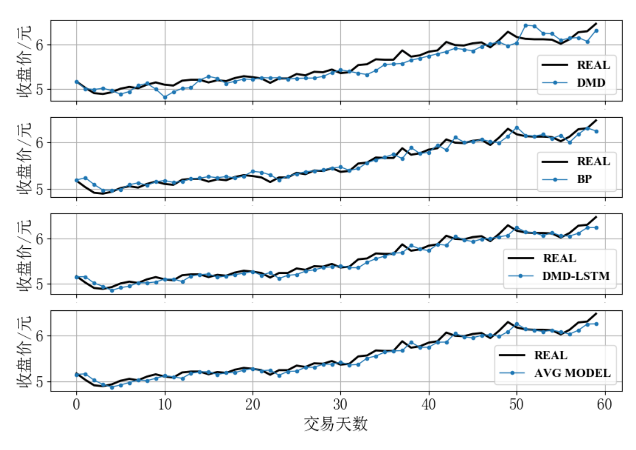


图 10 单边上涨形态下模型预测结果对比图

Fig. 10 Comparison results of various models in stronger market

表 1　单边上涨阶段各模型预测结果比较

Table 1　Comparing the predict results of various models in stronger market

预测模型	RMSE	MAE	R2	DA
SVR	0.1160	0.0879	0.9307	0.4576
BP	0.0921	0.0708	0.9564	0.4548
DMD	0.1389	0.1136	0.9006	0.5593 *
LSTM	0.0914	0.0702	0.9570	0.5198
DMD-LSTM	0.0898	0.0698	0.9592	0.4934
融合模型 *	0.0886 *	0.0686 *	0.9596 *	0.5127

股票价格单边下跌阶段，市场总体走弱，相关模型的预测结果如表 2 所示。类似于单边上涨形态的预测结果，DMD 在价格预测方面精度较差，但在预测价格趋势方向准确性依然好于 BP 和 SVR 两种基准模型。DMD-LSTM 在单模型上的预测精度和 LSTM 单模型相当，取得了较高的预测准确性，RMSE 误差比基准模型低了平均 17.17%。值得注意的是该模型在这个市场形态下的预测 DA 值相对较高，这与单边上涨情况的预测情况不同，可能原因是单边市场下的局部波动导致了这种不稳定的情况。在模型融合方面，预测结果进行加权平均后的预测误差 RMSE 在 DMD-LSTM 模型的基础上进一步下降了 3.72%，同时模型拟合优度最佳且 DA 值最大，实现了模型预测性能的最优。

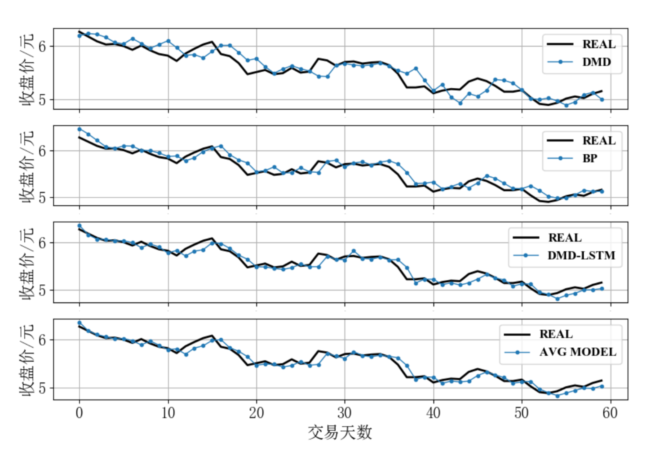


图 11　单边下跌形态下模型预测结果对比图

Fig. 11　Comparison results of various models in weaker market

表 2　单边下跌形态下各个模型预测结果对比

Table 2　Comparing the predict results of various models in weaker market

预测模型	RMSE	MAE	R2	DA
SVR	0.1094	0.0809	0.9104	0.4746
BP	0.1247	0.0968	0.8834	0.5051
DMD	0.1593	0.1268	0.8102	0.5424
LSTM	0.0984	0.0801	0.9275	0.5424
DMD-LSTM	0.0989	0.0798	0.9235	0.5953
融合模型 *	0.0970 *	0.0772 *	0.9294 *	0.5953 *

DMD 方法进行有效的模态特征提取的前提是市场趋势存在延续性，当市场进入震荡调整期，行业板块联动效应较弱，DMD 算法较难从短期的时间序列数据中分解出有效的趋势信息。

股票价格震荡阶段，价格波动变化较大，相关模型的预测结果如表 3 所示。由于预测结果滞后性严重，DMD 算法对于价格的拟合预测精度出现大幅下降，拟合优度为负数，表明无法进行有效的价格预测。相比于其他单模型，LSTM 模型的预测精度最高，由于行业板块下提取的趋势变化的信息不准确，DMD-LSTM 中的模态特征相当于引入外部噪声，因此其预测精度低于 LSTM 模型。进一步将预测结果平均加权进行模型融合后，尽管预测结果的 RMSE 和 R^2 都实现了小幅度提升，但不排除偶然因素导致的影响。实证结果表明，基于 DMD-LSTM 模型的股票价格预测方法在市场震荡的情况进行应用还存在一定的缺陷。

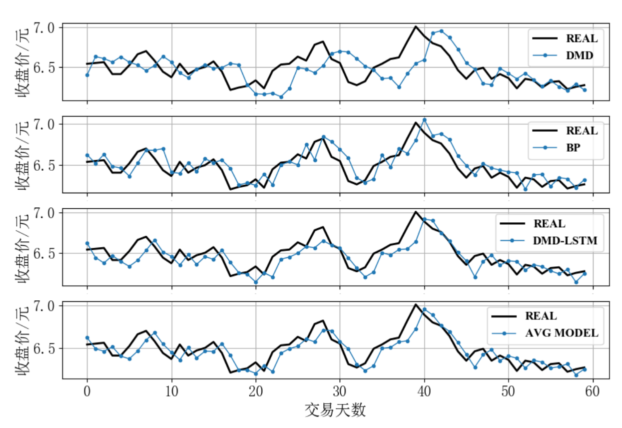


图 12　震荡阶段下模型预测结果对比图

Fig. 12　Comparison results of various models in shocking market

表 3　震荡阶段下各个模型预测结果对比

Table 3　Comparing the predict results of various models in shocking market

预测模型	RMSE	MAE	R2	DA
SVR	0.1222	0.0956	0.5561	0.4746
BP	0.1183	0.0934	0.5830	0.5627
DMD	0.2012	0.1600	-0.2033	0.5254
LSTM	0.1109	0.0908 *	0.6340	0.5720 *
DMD-LSTM	0.1150	0.0965	0.6059	0.4814
融合模型 *	0.1106 *	0.0930	0.6359 *	0.5322

4　结束语

本文提出的基于 DMD-LSTM 模型的股票价格预测方法，从国内股市相关行业板块存在的关联效应出发，通过 DMD 算法进行模态特征提取来获取行业板块间的潜在关联因素和趋势信息，然后将 LSTM 模型应用于股票价格预测，充分的利用了时序信息和模态特征。

通过鞍钢股份不同市场背景下的对比实验表明，相比于传统机器学习预测方法，本文模型在单边市场趋势显著的前提背景下，能取得更加准确的预测结果，包括更低的预测误差和更高的方向准确性。这表明在单边市场情况下，行业板块关联效应较为显著，利用 DMD 提取的模态特征能有效提升模型的预测性能。在震荡市场情况下，模型效果并不理想，其原因在于市场价格波动明显，价格变化影响因素复杂，行业板块联动效应较弱，相关模态特征提取方法很难获取有效的趋势信息。之后工作将进一步研究震荡市场情况下，板块关联效应同动态模态特征提取间的关系。

参考文献：

[1] Nazário R T F, Silva J L E, Sobreiro V A, *et al.* A literature review of technical analysis on stock markets [J]. Quarterly Review of Economics & Finance, 2017, 66(1): 115-126.

[2] Weng Bin, Ahmed M A, Megahed F M. Stock market one-day ahead movement prediction using disparate data sources [J]. Expert Systems with Applications, 2017, 79(2): 153-163.

[3] Bradley T E. The transmission of shocks among S&P indexes [J]. Applied Financial Economics, 2002, 12(4): 285-290.

[4] 周武. 股改以来我国股市行业板块间联动效应的实证研究[D]. 合肥: 安徽财经大学, 2012. (Zhou Wu. Co-movement in different industry index after stock reform in China [D]. Hefei: Anhui University of Finance and Economics, 2012.)

[5] 魏国林. 我国股市联动效应和龙头股研究 [D]. 成都: 西南财经大学, 2013. (Wei Guolin. The research of stock linkage effect and bellwether [D]. Chengdu: Southwestern University of Finance and Economics, 2013.)

[6] 温凯迪. 我国证券市场板块联动效应研究 [D]. 合肥: 安徽大学, 2013. (Wen Kaidi. Research on plate linkage phenomenon in China’s securities market [D]. Hefei: Anhui University, 2013.)

[7] Xu Mengjia, Shang Pengjian, Lin Aijing. Cross-correlation analysis of stock markets using EMD and EEMD [J]. Physical A Statistical Mechanics & Its Applications, 2016, 442 (1): 82-90.

[8] Mann J, Kutz J N. Dynamic mode decomposition for financial trading strategies [J]. Quantitative Finance, 2015, 2015(8): Article ID 1170194.

[9] 李晋. 基于小波分析和 GA-SVR 模型的股指期货价格预测方法研究 [D]. 广州:华南理工大学, 2011. (Research on stock index future price forecasting methods based on wavelet transform and GA-SVR model [D]. Guangzhou: South China University of Technology, 2011.)

[10] 郭志强, 曾亚丽, 杨杰, 等. 基于 MPCA-RBF 模型的证券市场指数时间序列预测 [J]. 计算机应用研究, 2017, 34(11): 3299-3302. (Guo Zhiqiang, Zeng Yali, Yang Jie, *et al.* Time series forecasting of stock market index based on MPCA-RBF model [J]. Journal of Application Research of Computers, 2017, 34(11): 3299-3302.)

[11] 史文静, 高岩. EMD 结合 RBF 神经网络新混合模型及股指期货价格预测 [J]. 经济数学, 2015(1): 47-51. (Shi Wenjing, Gao Yan. Prediction of the Chinese stock index futures market based on new EMD-RBF model [J]. Journal of Quantitative Economics, 2015(1): 47-51.)

[12] Schmid P J. Dynamic Mode Decomposition of numerical and experimental data [J]. Journal of Fluid Mechanics, 2010, 656 (10): 5-28.

[13] Cui Lingxiao, Long Wen. Trading strategy based on dynamic mode decomposition: Tested in Chinese stock market [J]. Physica A Statistical Mechanics & Its Applications, 2016, 461(1): 498-508.

[14] 龙奥明, 毕秀春, 张曙光. 基于 LSTM 神经网络的黑色金属期货套利策略模型 [J]. 中国科学技术大学学报, 2018, 48(2): 125-132. (Long Aoming, Bi Xiuchun, Zhang Shuguang. An arbitrage strategy model for ferrous metal futures based on LSTM neural network [J]. Journal of University of Science and Technology of China, 2018, 48(2): 125-132.)

[15] Graves A. Supervised Sequence Labelling [M]//Supervised Sequence Labelling with Recurrent Neural Networks. Berlin: Springer, 2012: 5-13.